



TITLE:

良条件行列近似問題に対する逐次射影法 (最適化の基礎理論と応用)

AUTHOR(S):

田中, 未来; 中田, 和秀

CITATION:

田中, 未来 ...[et al]. 良条件行列近似問題に対する逐次射影法 (最適化の基礎理論と応用). 数理解析研究所講究録 2014, 1879: 34-43

ISSUE DATE:

2014-04

URL:

<http://hdl.handle.net/2433/195626>

RIGHT:

良条件行列近似問題に対する逐次射影法

田中未来*

中田和秀†

概要

本論文では, 条件数制約と符号制約を満たす正定値行列あるいは相関行列の中で与えられた行列に最も近いものを求める問題を考える. これらの問題は条件数制約がなす閉凸錐とそれ以外の制約がなす凸多面集合の共通部分への射影とみなすことができる. そのため, この問題は複数の凸集合の共通部分への射影を求めるためのアルゴリズムである逐次射影法で解くことができる. 簡単な収束性の証明を与えた上で, 計算実験によりこの方法の有効性を確認する.

1 良条件行列近似問題

行列の条件数 (最大特異値と最小特異値の比) が 1 に近いときその行列を良条件であるといい, そうでないとき悪条件であるという. 行列が良条件であることは工学の多くの分野において望ましい性質である. 以下にその例を示す:

例 1 (線形方程式系の誤差解析). 線形方程式系 $Ax = b$ の右辺ベクトルの誤差が解に及ぼす影響を考える. $Ax = b$ の解を x^* とし, 右辺を $b + \Delta b$ に摂動させた線形方程式形の解を $x^* + \Delta x^*$ とする. このとき Δx^* は右辺ベクトルの摂動が解に与える影響とみなすことができる. この大きさは係数行列 A の条件数 $\text{cond}(A)$ を用いて次のように上から抑えることができる:

$$\frac{\|\Delta x^*\|_2}{\|x^*\|_2} \leq \text{cond}(A) \frac{\|\Delta b\|_2}{\|b\|_2}.$$

係数行列 A をも摂動させた場合も同様の解析ができる. 詳細は Horn and Johnson [11, Section 5.8] などを参照されたい.

例 2 (最急降下法の収束解析). 無制約最適化問題 $\min f(x)$ を最急降下法で解くときの漸近的な収束率を考える. 最急降下法によって生成された点列 $\{x^{(k)}\}$ が $k \rightarrow \infty$ のとき局所的最適解 x^* に収束したとする. このときの収束の次数は 1 次であり, 漸近的な収束比は

$$\left(\frac{1 - \text{cond}(\nabla^2 f(x^*))}{1 + \text{cond}(\nabla^2 f(x^*))} \right)^2$$

* 東京工業大学 大学院社会理工学研究科 経営工学専攻, ☎ 152-8552 東京都目黒区大岡山 2-12-1-W9-60, tanaka.m.aa@m.titech.ac.jp

† 東京工業大学 大学院社会理工学研究科 経営工学専攻, ☎ 152-8552 東京都目黒区大岡山 2-12-1-W9-60, nakata.k.ac@m.titech.ac.jp

となることが知られている。証明は Luenberger [13, Section 12.5]などを参照されたい。

信号処理の分野では、条件数制約の下で分散共分散行列を最尤推定する方法が研究されている [1, 18]。具体的には、多変量正規分布を仮定したときに条件数制約の下で対数尤度を最大化する問題が1変数の最適化問題に帰着できることが示されている。しかしながら、他の分布を仮定する場合や尤度の最大化でない場合にこのようなことができるとは限らない。そのような場合、与えられた行列に最も近い良条件な行列を求めるということは自然な発想だろう。Tanaka and Nakata [16]はこの問題を次のような n 次対称行列の空間 \mathcal{S}^n 上の最適化問題として定式化している：

$$\begin{array}{ll} \text{minimize} & \|X - \widehat{X}\| \\ \text{subject to} & X \in \mathcal{S}_+^n, \\ & \text{cond}(X) \leq \kappa. \end{array} \quad (1)$$

ここで $\widehat{X} \in \mathcal{S}^n$ は与えられた行列、 $X \in \mathcal{S}^n$ は変数、 $\mathcal{S}_+^n \subset \mathcal{S}^n$ は n 次半正定値行列がなす錐、 $\kappa \geq 1$ は X の条件数の上限を定める定数である。なお、本論文では $\text{cond}(O) = 1$ と定義する。彼らは問題 (1) が固有値分解程度の計算時間で解けることを示している。

金融工学の世界では、分散共分散行列のいくつかの成分がアナリストの直観に反する場合、それらの成分を事後的に 0 にするなどの補正を行うことがある [3, Section 10.2]。しかしながら、このような補正を施した行列は良条件であることはおろか半正定値になる保証すらない。したがって、このような場合、条件数を考慮して補正を行なう必要がある。このような符号制約を満たす良条件な行列で与えられた行列を近似する問題は次のように定式化できる：

$$\begin{array}{ll} \text{minimize} & \|X - \widehat{X}\| \\ \text{subject to} & X \in \mathcal{S}_+^n, \\ & \text{cond}(X) \leq \kappa, \\ & X_{ij} \geq 0 \quad ((i, j) \in P), \\ & X_{ij} \leq 0 \quad ((i, j) \in N). \end{array} \quad (2)$$

ここで $P, N \subset \{1, \dots, n\}^2$ は符号制約を課す添字の集合である。

この問題は共分散選択 [4, 5, 8] と関連している。共分散行列が Σ であるような多変量正規分布について、変数 i と変数 j が条件付き独立であるときかつそのときに限り $[\Sigma^{-1}]_{ij} = 0$ となることが知られている。事前に条件付き独立であることがわかっている変数の組があったとき、問題 (1) を解くことで標本共分散行列 $\widehat{\Sigma}$ を条件付き独立性を満たすように補正することができる。具体的には、条件付き独立な変数の組の集合を I_0 とおき、 $\widehat{X} = \widehat{\Sigma}^{-1}$ 、 $\kappa = c$ 、 $P = N = I_0$ とした問題 (1) を解く。得られた最適解 X^* について $\Sigma = (X^*)^{-1}$ とすれば Σ は $\text{cond}(\Sigma) \leq c$ と任意の $(i, j) \in I_0$ に対して $[\Sigma^{-1}]_{ij} = 0$ を満たす $\widehat{\Sigma}$ に近い行列になっている。

相関行列に対して同様の補正を行ないたい場合は、問題 (1) に対角成分を 1 に固定する制約を追

加した以下の問題を解けばよい:

$$\begin{array}{ll}
 \text{minimize} & \|X - \widehat{X}\| \\
 \text{subject to} & X \in S_+^n, \\
 & \text{cond}(X) \leq \kappa, \\
 & X_{ii} = 1 \quad (i = 1, \dots, n), \\
 & X_{ij} \geq 0 \quad ((i, j) \in P), \\
 & X_{ij} \leq 0 \quad ((i, j) \in N).
 \end{array} \tag{3}$$

この問題は $P = N = \emptyset$ のとき, 条件数制約のついた近似相関行列問題となる. 近似相関行列問題に関する研究は Higham [10], Qi and Sun [14] などがあるが, 条件数制約を考慮したものは今のところ存在しない.

問題 (1) と問題 (1) は目的関数のノルムを適当に設定すれば対称錐最適化問題として表現できる. したがって, これらの問題は内点法を用いて多項式時間で解くことができる. しかしながら, 等価な対称錐最適化問題が非常に大規模になるため, 大規模な問題例をこの方法で解くことは難しい.

本論文では, 問題 (1) と問題 (1) の目的関数において Frobenius ノルムを用いた問題に対する Tanaka and Nakata [17] によって提案された解法を紹介する. 具体的には, これらの問題を複数の凸集合の共通部分への直交射影を求める問題とみなし, 逐次射影法 [2, 7, 9] を適用する. このアルゴリズムの内部では問題 (1) を Tanaka and Nakata [16] による 2 分探索で解く.

本論文の構成は以下の通りである: 2 節では, 一般的な逐次射影法の説明を行なった後, Frobenius ノルムを用いた問題 (1) と問題 (1) に対して逐次射影法を適用することを考える. ここでは, これらの問題に対して逐次射影法が大域的収束することと, 逐次射影法の各反復が簡単に計算できることを示す. その後, 3 節で逐次射影法の有効性を示す計算実験の結果を示し, 4 節で結論および今後の課題を述べる.

なお, 以下ではすべての i に対して $(i, i) \notin N$ であることを仮定する. これは, もしそうでないとすると, 問題 (1) は $X = O$ 以外の実行可能解を持たず, 問題 (1) は実行不能となるためである.

2 逐次射影法

逐次射影法を導入するために, まず次の問題を考えよう:

$$\begin{array}{ll}
 \text{minimize} & \|x - \widehat{x}\| \\
 \text{subject to} & x \in C_i \quad (i = 1, \dots, m),
 \end{array} \tag{4}$$

ここで $\widehat{x} \in \mathbb{R}^n$ は与えられたベクトル, $x \in \mathbb{R}^n$ は決定変数, $\|\cdot\|$ は \mathbb{R}^n 上の内積から誘導されたノルム, C_1, \dots, C_m は \mathbb{R}^n 上の閉凸集合である. この問題の最適解は \widehat{x} の $\bigcap_{i=1}^m C_i$ への直交射影である.

逐次射影法は問題 (2) を解くための古典的なアルゴリズムである. 具体的には各閉凸集合 C_i に逐次的に直交射影を繰り返すことで最適解に収束する点列を生成する. このアルゴリズムの擬似コードはアルゴリズム 1 に示すように極めて単純である. なお, $P_i(\widehat{x})$ は \widehat{x} の各閉凸集合 C_i に対する直

交射影, すなわち, 次の問題の最適解を意味する:

$$\begin{array}{ll} \text{minimize} & \|x - \widehat{x}\| \\ \text{subject to} & x \in C_i. \end{array}$$

アルゴリズム 1 問題 (2) に対する逐次射影法

```

1:  $x_m^{(0)} := \widehat{x}, y_1^{(0)}, \dots, y_m^{(0)} := 0$ , and  $k = 0$ 
2: while  $x_m^{(k)}$  が近似最適解でない do
3:    $x_0^{(k+1)} := x_m^{(k)}$ 
4:   for  $i = 1, \dots, m$  do
5:      $x_i^{(k+1)} := P_i(x_{i-1}^{(k+1)} + y_i^{(k)})$ 
6:      $y_i^{(k+1)} := x_{i-1}^{(k+1)} + y_i^{(k)} - x_i^{(k+1)}$ 
7:   end for
8:    $k := k + 1$ 
9: end while

```

このアルゴリズムは Dykstra [7] によって複数の閉凸錐に対して提案され, その後 Boyle and Dykstra [2] によって Hilbert 空間上の閉凸集合に対して拡張された. そのため, このアルゴリズムは Dykstra のアルゴリズムとも呼ばれる. また, Han [9] はこの拡張されたアルゴリズムについて Euclid 空間上で精密に解析を行なった. 次の定理はこのアルゴリズムの大域的収束性を意味するものである.

定理 3 (Han [9, Theorem 4.8]). C_1, \dots, C_p を凸多面集合, C_{p+1}, \dots, C_m を閉凸集合とし, $(\bigcap_{i=1}^p C_i) \cap (\bigcap_{i=p+1}^m \text{int } C_i) \neq \emptyset$ を満たすものとする. このとき, アルゴリズム 1 によって生成される点列 $\{x_m^{(k)}\}$ は $k \rightarrow \infty$ のとき問題 (2) の最適解に収束する.

このアルゴリズムを問題 (1) と問題 (1) に適用することを考える. まず, 問題 (1) に対して次のような集合を定める:

$$\begin{aligned} C_{\text{poly}} &= \{X \in \mathcal{S}^n : X_{ij} \geq 0 \ ((i, j) \in P), X_{ij} \leq 0 \ ((i, j) \in N)\}, \\ C_{\text{cond}} &= \{X \in \mathcal{S}_+^n : \text{cond}(X) \leq \kappa\}. \end{aligned}$$

また, 問題 (1) に対しては C_{poly} を次のように書き換える:

$$C_{\text{poly}} = \{X \in \mathcal{S}^n : X_{ii} = 1 \ (i = 1, \dots, n), X_{ij} \geq 0 \ ((i, j) \in P), X_{ij} \leq 0 \ ((i, j) \in N)\}.$$

どちらの問題に対しても C_{poly} が閉凸集合であることは明らかである. 次の補題では C_{cond} が閉凸集合であることを示す.

補題 4. C_{cond} は閉凸集合.

Proof. まず, C_{cond} が閉であることを示す. $\text{cond}(O) = 1$ と定義したことを思い出すと, 次が成り立つ:

$$C_{\text{cond}} = \{X \in S_+^n : \lambda_{\max}(X) \leq \kappa \lambda_{\min}(X)\}.$$

ここで, $\lambda_{\min}(X)$ と $\lambda_{\max}(X)$ は X の最小固有値と最大固有値を表す. $\lambda_{\min}(\cdot)$ と $\lambda_{\max}(\cdot)$ の連続性 [11, Theorem 2.4.9.2] より, C_{cond} は閉.

次に, C_{cond} が凸であることを示す. そのために, 任意に $A, B \in C_{\text{cond}}$ と $\alpha \in [0, 1]$ をとる. さらに, $C = (1 - \alpha)A + \alpha B$ とし, v_{\min} と v_{\max} をそれぞれ $\lambda_{\min}(C)$ と $\lambda_{\max}(C)$ に対応する正規化した固有ベクトルとする. このとき, 次の不等式が成り立つ:

$$\lambda_{\min}(C) = v_{\min}^T C v_{\min} = (1 - \alpha) v_{\min}^T A v_{\min} + \alpha v_{\min}^T B v_{\min} \geq (1 - \alpha) \lambda_{\min}(A) + \alpha \lambda_{\min}(B).$$

同様に, 次の不等式を示すことができる:

$$\lambda_{\max}(C) \leq (1 - \alpha) \lambda_{\max}(A) + \alpha \lambda_{\max}(B).$$

$\lambda_{\max}(A) \leq \kappa \lambda_{\min}(A)$ と $\lambda_{\max}(B) \leq \kappa \lambda_{\min}(B)$ より, 次が成り立つ:

$$\text{cond}(C) = \frac{\lambda_{\max}(C)}{\lambda_{\min}(C)} \leq \frac{(1 - \alpha) \lambda_{\max}(A) + \alpha \lambda_{\max}(B)}{(1 - \alpha) \lambda_{\min}(A) + \alpha \lambda_{\min}(B)} \leq \kappa.$$

したがって, C_{cond} は凸. □

アルゴリズム 1 の各反復の計算は直交射影 $P_i(\cdot)$ の計算が簡単であれば高速に実行できる. 以下に示すように, 問題 (1) と問題 (1) で Frobenius ノルムを用いた場合は直交射影の計算を高速に実行できる. $P_{\text{poly}}(\cdot)$ と $P_{\text{cond}}(\cdot)$ をそれぞれ C_{poly} と C_{cond} への射影としよう. まず, $P_{\text{poly}}(\cdot)$ は要素ごとに計算を分解できるので, 簡単に計算できる. 実際, 問題 (1) における $P_{\text{poly}}(\cdot)$ は次のようになる:

$$[P_{\text{poly}}(X)]_{ij} = \begin{cases} 1 & i = j \text{ のとき,} \\ 0 & i \neq j \text{ で符号制約を違反しているとき,} \\ X_{ij} & \text{それ以外.} \end{cases}$$

また, $P_{\text{cond}}(\cdot)$ の計算は Tanaka and Nakata [16] の 2 分探索を用いれば固有値分解程度の計算量で実行可能である. 最後に, 問題 (1) と問題 (1) に対する逐次射影法をアルゴリズム 2 に示す.

次の定理は問題 (1) と問題 (1) に対するこのアルゴリズムの大域的収束性を示すものである.

定理 5. $\kappa > 1$ のとき, アルゴリズム 2 によって生成される点列 $\{X_{\text{poly}}^{(k)}\}$ は $k \rightarrow \infty$ のとき問題 (1) あるいは問題 (1) の最適解に収束する.

証明. 明らかに C_{poly} は凸多面集合. また, 補題 4 より, C_{cond} は凸. さらに, $\kappa > 1$ のとき $I \in C_{\text{poly}} \cap \text{int} C_{\text{cond}}$ が成り立つ. したがって, 定理 3 が適用できるため, 点列 $\{X_{\text{poly}}^{(k)}\}$ は最適化に収束することがわかる. □

注意 6. ここでの仮定 $\kappa > 1$ は本質的ではない. なぜならば, $\kappa = 1$ のときは C_{cond} が凸多面集合 $\{\alpha I : \alpha \geq 0\}$ となるため, $I \in C_{\text{poly}} \cap C_{\text{cond}}$ より, 定理 3 を適用できる. しかしながら, この場合は 2 つの問題の最適解がともに $(\text{tr}(\hat{X})/n)I$ となるため, 逐次射影法を適用する意味がない.

アルゴリズム 2 問題 (1) と問題 (1) に対する逐次射影法

```

1:  $Y_{\text{poly}}^{(0)}, Y_{\text{cond}}^{(0)} := O, X_{\text{poly}}^{(1)} := P_{\text{poly}}(\widehat{X}), k = 1$ 
2: while  $X_{\text{poly}}^{(k)}$  が近似最適解でない do
3:    $X_{\text{cond}}^{(k+1)} := P_{\text{cond}}(X_{\text{poly}}^{(k)} + Y_{\text{cond}}^{(k)})$ 
4:    $Y_{\text{cond}}^{(k+1)} := X_{\text{poly}}^{(k)} + Y_{\text{cond}}^{(k)} - X_{\text{cond}}^{(k+1)}$ 
5:    $X_{\text{poly}}^{(k+1)} := P_{\text{poly}}(X_{\text{cond}}^{(k+1)} + Y_{\text{poly}}^{(k)})$ 
6:    $Y_{\text{poly}}^{(k+1)} := X_{\text{cond}}^{(k+1)} + Y_{\text{poly}}^{(k)} - X_{\text{poly}}^{(k+1)}$ 
7:    $k := k + 1$ 
8: end while

```

注意 7. アルゴリズム 2 は次のような問題に対しても適用できる:

$$\begin{array}{ll}
 \text{minimize} & \|X - \widehat{X}\| \\
 \text{subject to} & \mathcal{A}X = b, \\
 & X \in S_+^n, \\
 & \text{cond}(X) \leq \kappa.
 \end{array}$$

ここで $\mathcal{A}: S^n \rightarrow \mathbb{R}^m$ は線形写像, $b \in \mathbb{R}^m$, $\widehat{X} \in S^n$ は定数, $X \in S^n$ は決定変数である. ここで

$$C_{\text{poly}} = \{X \in S^n : \mathcal{A}X = b\}$$

とし, $\mathcal{A}^\top: \mathbb{R}^m \rightarrow S^n$ を \mathcal{A} の随伴とすると, この凸多面集合への射影は次のように書ける:

$$P_{\text{poly}}(\widehat{X}) = \widehat{X} + \mathcal{A}^\top (\mathcal{A}\mathcal{A}^\top)^{-1} (b - \mathcal{A}\widehat{X}).$$

したがって, m が小さかったり, $\mathcal{A}\mathcal{A}^\top$ の Cholesky 分解や $(\mathcal{A}\mathcal{A}^\top)^{-1}$ が疎であったりするとき, $P_{\text{poly}}(\cdot)$ の計算量は小さくなるため, アルゴリズム 2 の各反復は高速に計算できる.

3 計算実験

逐次射影法の有効性を確認するために行なった計算実験の結果を示す. ここでは, 逐次射影法と内点法を用いて Frobenius ノルムを用いた問題 (1) と問題 (1) の複数の問題例を解いた. 逐次射影法は MATLAB 7.14 を用いて実装し, 終了条件は $X_{\text{poly}}^{(k)} \in S_+^n$ と $\text{cond}(X_{\text{poly}}^{(k)}) \leq (1 + 10^{-6})\kappa$ を同時に満たすこととした. 比較のために, Frobenius ノルムを用いた問題 (1) と問題 (1) をモデリング言語 YALMIP 3 [12] を用いてモデリングし, 得られた対称錐最適化問題を内点法が実装された対称錐最適化ソルバ SeDuMi 1.3 [15] を用いて解いた.

問題例は以下のように生成した. まず, $[-1, +1]$ 上の一様分布に従う擬似乱数を用いて行列 $U \in \mathbb{R}^{n \times n}$ の各要素を生成し, $\widehat{X} = U + U^\top$ とした. さらに, P と N を \widehat{X} の各成分のうち大きい方と小さい方から順に $2n$ 個の要素に対応する添字集合とした. また, $\kappa = 10^6$ とした.

問題 (1) と問題 (1) の小規模な問題例と解いた結果を表 1 と表 2 に示す。“ n ” の列は行列のサイズを, “計算時間 [秒]” の列は逐次射影法と内点法それぞれの計算時間を, “反復回数” の列は逐次射影法の反復回数を表す。これらの表から, 問題のサイズが大きくなるにつれて内点法の計算時間が急激に増加する一方で, 逐次射影法の計算時間の増加は緩やかであることが見て取れる。

表 1 問題 (1) の小規模な問題例に対する内点法と逐次射影法の結果

n	内点法	逐次射影法	
	計算時間 [秒]	計算時間 [秒]	反復回数
20	3.5×10^{-1}	5.6×10^{-2}	98
30	1.1×10^0	9.7×10^{-2}	133
40	3.1×10^0	9.7×10^{-2}	94
50	9.1×10^0	1.1×10^{-1}	75
60	$2.6 \times 10^{+1}$	1.3×10^{-1}	79
70	$1.0 \times 10^{+2}$	1.8×10^{-1}	78
80	$3.1 \times 10^{+2}$	1.7×10^{-1}	70
90	$7.4 \times 10^{+2}$	2.3×10^{-1}	69
100	$1.6 \times 10^{+3}$	3.0×10^{-1}	71

表 2 問題 (1) の小規模な問題例に対する内点法と逐次射影法の結果

n	内点法	逐次射影法	
	計算時間 [秒]	計算時間 [秒]	反復回数
20	3.7×10^{-1}	5.5×10^{-2}	97
30	9.7×10^{-1}	9.9×10^{-2}	130
40	3.3×10^0	1.7×10^{-1}	154
50	8.8×10^0	2.3×10^{-1}	176
60	$2.7 \times 10^{+1}$	2.9×10^{-1}	186
70	$1.0 \times 10^{+2}$	3.7×10^{-1}	195
80	$3.1 \times 10^{+2}$	4.8×10^{-1}	211
90	$7.9 \times 10^{+2}$	7.1×10^{-1}	219
100	$1.4 \times 10^{+3}$	9.8×10^{-1}	238

次に, 逐次射影法を用いて大規模な問題を解いた結果を表 3 と表 4 に示す。これらの表から内点法では現実的な計算時間で解くことができないであろう問題例を逐次射影法を用いることで高速に解くことができることがわかる。また, 問題 (1) の場合は行列のサイズが反復回数にほとんど依存しないのに対し, 問題 (1) の場合は行列のサイズの増加に伴って反復回数が増加することがわかる。

表 3 問題 (1) の大規模な問題例に対する逐次射影法の結果

n	計算時間 [秒]	反復回数
128	4.3×10^{-1}	60
256	1.4×10^0	49
512	5.1×10^0	45
1024	$2.4 \times 10^{+1}$	42
2048	$1.5 \times 10^{+2}$	38
4096	$1.1 \times 10^{+3}$	36
8192	$8.4 \times 10^{+3}$	35

表 4 問題 (1) の大規模な問題例に対する逐次射影法の結果

n	計算時間 [秒]	反復回数
128	1.1×10^0	237
256	6.0×10^0	308
512	$4.5 \times 10^{+1}$	402
1024	$2.7 \times 10^{+2}$	481
2048	$2.5 \times 10^{+3}$	654
4096	$2.5 \times 10^{+4}$	826
8192	$2.7 \times 10^{+5}$	1115

最後に, 逐次射影法が生成した点列 $\{X_{\text{poly}}^{(k)}\}$ が最終的な解 X^* に収束する様子を見てみよう。問題 (1) の $n = 100$ の問題例を逐次射影法で解いて得られた点列 $\{X_{\text{poly}}^{(k)}\}$ について, 反復回数を横軸にとって誤差ノルム $\|X_{\text{poly}}^{(k)} - X^*\|$ をプロットしたものを図 1 に示す。この図から $\{X_{\text{poly}}^{(k)}\}$ が X^* に 1 次収束する様子が見て取れる。問題 (1) の $n = 100$ の問題例についても同様の結果が得られた (図 2)。他の問題例についても同様の結果を得た。

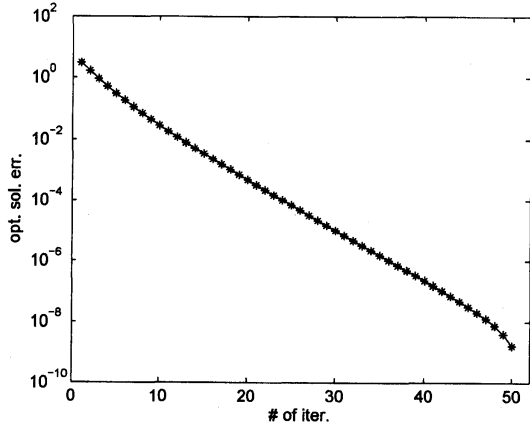


図1 問題(1)における誤差ノルム $\|X_{\text{poly}}^{(k)} - X^*\|_F$ の挙動

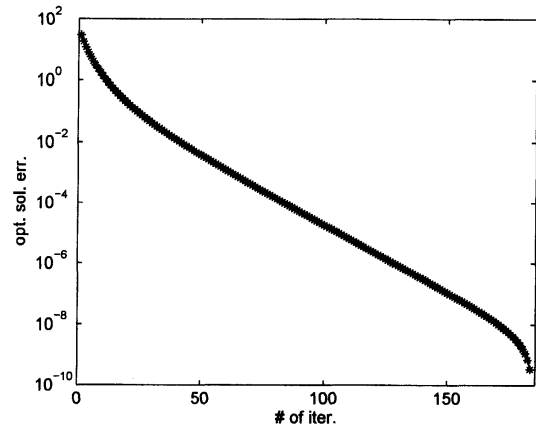


図2 問題(1)における誤差ノルム $\|X_{\text{poly}}^{(k)} - X^*\|_F$ の挙動

4 結論と今後の課題

本論文では Tanaka and Nakata [17] の良条件行列近似問題に対する逐次射影法を紹介した。このアルゴリズムは良条件行列がなす閉凸錐と符号制約などの線形制約がなす凸多面集合への射影を逐次的に繰り返すものである。このアルゴリズムには大きく2つの長所がある。1つはアルゴリズムが単純なため実装が簡単であることであり、もう1つは計算実験の結果で見たように実用的に1次収束を見込むことができる点である。

この1次収束性を理論的に証明することは今後の大きな課題である。Higham [10] は近似相関行列問題に対する逐次射影法が最良の場合に1次収束すると述べており、最悪の場合の収束性の解析は依然として行われていない。また、Deutsch and Hundal [6] はすべての閉凸集合が部分空間である場合の1次収束性を証明している。しかしながら、このアルゴリズムが1次収束するための必要十分条件は未だに知られていない。

また、今回は Frobenius ノルムを用いた場合に逐次射影法が適用できることを示したが、他のノルムを用いた場合に対して効率的なアルゴリズムを設計することも重要な今後の課題である。例えば、共分散行列を推定する場合、成分によって推定値の確からしさが異なる場合がある。このような場合、(重みづけをしていない) Frobenius ノルム $\|X\|_F = \sqrt{\sum_{i,j} X_{ij}^2}$ を用いるよりも、重みづけをした $\|H \circ X\|_F = \sqrt{\sum_{i,j} H_{ij}^2 X_{ij}^2}$ というようなノルムを用いることで、確からしさを考慮して行列の補正を行なう方が望ましい。しかしながら、このノルムはユニタリ相似不変でないため、 $P_{\text{cond}}(\cdot)$ の計算において Tanaka and Nakata [16] の2分探索法を直接適用することができない。

参考文献

- [1] A. AUBRY, A. DE MAIO, L. PALLOTTA, AND A. FARINA: Maximum likelihood estimation of a structured covariance matrix with a condition number constraint, *IEEE Transaction on Signal Processing* **60** (2012) 3004–3021.
- [2] J. P. BOYLE AND R. L. DYKSTRA: A method for finding projections onto the intersection of convex sets in Hilbert spaces, in *Advances in Order Restricted Statistical Inference*, R. Dykstra, T. Robertson, and F. T. Wright (eds.), 1985, pp. 28–47.
- [3] G. CORNUEJOLS AND R. TÜTÜNCÜ: *Optimization Methods in Finance*, Cambridge University Press, Cambridge, 2007.
- [4] A. D’ASPREMONT, O. BANERJEE, AND L. EL GHAOU: First-order methods for sparse covariance selection, *SIAM Journal on Matrix Analysis and Applications* **30** (2008) 56–66.
- [5] A. P. DEMPSTER: Covariance selection, *Biometrics* **28** (1972) 157–175.
- [6] F. DEUTSCH AND H. HUNDAL: The rate of convergence for the method of alternating projections II, *Journal Mathematical Analysis and Applications* **205** (1997) 381–405.
- [7] R. L. DYKSTRA: An algorithm for restricted least squares regressions, *Journal of the American Statistical Association* **78** (1983) 837–842.
- [8] J. FRIEDMAN, T. HASTIE, AND R. TIBSHIRANI: Sparse inverse covariance estimation with the graphical lasso, *Biostatistics* **9** (2008) 432–441.
- [9] S.-P. HAN: A successive projection method, *Mathematical Programming* **40** (1988) 1–14.
- [10] N. J. HIGHAM: Computing the nearest correlation matrix—a problem from finance, *IMA Journal of Numerical Analysis* **22** (2002) 392–343.
- [11] R. A. HORN AND C. R. JOHNSON: *Matrix Analysis (Second Edition)*, Cambridge University Press, 2012.
- [12] J. LÖFBERG: YALMIP: A toolbox for modeling and optimization in MATLAB, in *Proceedings of the CACSD Conference*, 2004, pp. 284–289.
- [13] D. G. LUENBERGER AND Y. YE: *Linear and Nonlinear Programming (Third Edition)*, Springer, New York, 2010.
- [14] H. QI AND D. SUN: A quadratically convergent newton method for computing the nearest correlation matrix, *SIAM Journal on Matrix Analysis and Applications* **28** (2006) 360–385.
- [15] J. F. STURM: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optimization Methods and Software* **11** (1999) 625–653.
- [16] M. TANAKA AND K. NAKATA: Positive definite matrix approximation with condition number constraint, *Optimization Letters*, 2013, Online First (DOI: 10.1007/s11590-013-0632-7).
- [17] M. TANAKA AND K. NAKATA: Successive projection method for well-conditioned matrix

approximation problems, *Technical Report 2013-12, Department of Industrial Engineering and Management, Tokyo Institute of Technology*, 2013.

- [18] J. H. WON AND S.-J. KIM: Maximum likelihood covariance estimation with a condition number constraint, in *Proceedings 40th Asilomar Conference Signals, Systems, and Computers*, 2006, pp. 1445–1449.